Enlargement of Filtrations — A Primer

P. Ouwehand

Department of Mathematical Sciences
Stellenbosch University

Contents

1 Introduction **1**

2 Enlargements of Filtrations via Changes of Measure **4**

3 Doob–Meyer Decompositions via Girsanov's Theorem **8**

4 Jacod's Criterion for Initial Enlargements **9**

5 Stochastic Integrals under Enlargements **14**

6 Initial Enlargements in the Brownian World **14**

A Monotone Class Theorems **19**

B Completions, the Usual Hypotheses, etc. **22**

C Regular Conditional Probabilities **23**

1 Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with two filtrations $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$, both satisfying the usual conditions, and such that $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all $t \geq 0$.

We will study the following hypothesis:

$$\boxed{\text{Hypothesis } \mathbf{H}'\text{: Every } (\mathbb{F},\mathbb{P})\text{–semimartingale is a } (\mathbb{G},\mathbb{P})\text{–semimartingale.}}$$

Note that by the Doob–Meyer Decomposition Theorem and the Fundamental Theorem of Local Martingales, $(\mathbf{H}')$ is equivalent to each of the following:

(i) Every $(\mathbb{F}, \mathbb{P})$–local martingale is a $(\mathbb{G}, \mathbb{P})$–semimartingale.

(ii) Every bounded $(\mathbb{F}, \mathbb{P})$–martingale is a $(\mathbb{G}, \mathbb{P})$–semimartingale.

Example 1.1 Let W be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, with natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, and expand $\mathbb{F}$ by adding knowledge of W_T for some $T > 0$, i.e. consider the enlarged filtration $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ which is the smallest filtration satisfying the usual conditions

such that $\sigma(W_T) \cup \mathcal{F}_t \subseteq \mathcal{H}_t$ for all $t \geq 0$. We will show (following Protter[?]) that W is a $\mathbb{H}$–semimartingale (though no longer a $\mathbb{H}$–martingale) with decomposition

$$W_t = M_t + A_t \qquad \text{where} \quad M_t = W_t - \int_0^{t \wedge T} \frac{W_T - W_s}{T - s}\, ds$$

and M is a $\mathbb{H}$–martingale.

Without loss of generality, we may assume that $T = 1$. To show M is a $\mathbb{H}$–martingale, we will need to calculate $\mathbb{E}[W_t - W_s | \mathcal{H}_s]$. Since $W_t - W_s$ is independent of $\mathcal{F}_s$ for $t > s$, and since $\mathcal{H}_s$ is (essentially) $\mathcal{F}_s \vee \sigma(W_1 - W_s)$ it follows that $\mathbb{E}[W_t - W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s|W_1 - W_s]$. Symmetry now suggests that

$$\mathbb{E}[W_t - W_s|W_1 - W_s] = \frac{t-s}{1-s}(W_1 - W_s)$$

To prove this relation, let $0 \leq s < t \leq 1$ be rational numbers. Then there exists $n, j, k \in \mathbb{N}$ such that $s = \frac{j}{n}, t = \frac{k}{n}$. For $i = 0, 1, \ldots, n-1$, define $Y_i = W_{\frac{i+1}{n}} - W_{\frac{i}{n}}$. Then $W_1 - W_s = \sum_{i=j}^{n-1} Y_i, W_t - W_s = \sum_{i=j}^{k-1} Y_i$. Now as the Y_i are independent and identically distributed, symmetry dictates that

$$\mathbb{E}[W_t - W_s|W_1 - W_s] = \mathbb{E}\left[\sum_{i=j}^{k-1} Y_i \middle| \sum_{i=j}^{n-1} Y_i\right] = \frac{k-1}{n-j}\sum_{i=j}^{n-1} Y_i = \frac{t-s}{1-s}(W_1 - W_s)$$

Now as $W_t - W_s$ is independent of $\mathcal{F}_s$, we have

$$\mathbb{E}[W_t - W_s|\mathcal{H}_s] = \mathbb{E}[W_t - W_s|W_1 - W_s] = \frac{t-s}{1-s}(W_1 - W_s)$$

for all rationals $0 \leq s < t \leq 1$. Since W is continuous, a little argument shows that this relation holds for all $0 \leq s < t \leq 1$. Now, with M_t as defined above, and applying Fubini's Theorem (for conditional expectations) we have

$$\begin{aligned}
\mathbb{E}[M_t - M_s|\mathcal{H}_s] &= \mathbb{E}[W_t - W_s|\mathcal{H}_s] - \int_s^t \frac{\mathbb{E}[W_1 - W_u|\mathcal{H}_s]}{1-u}\, du \\
&= \frac{t-s}{1-s}(W_1 - W_s) - \int_s^t \frac{(1-u)(W_1 - W_s)}{(1-u)(1-s)}\, ds \\
&= 0
\end{aligned}$$

for $0 \leq s < t < 1$, i.e. M is a $\mathbb{H}$–martingale on $[0, 1)$. We want to show it that M is a martingale on all of $\mathbb{R}^+$. There may be a problem at $t = 1$ because of the possibility of explosions: We require that $\int_s^1 \frac{W_1 - W_s}{1-s}\, ds$ remains finite a.s. This will be guaranteed if we can show that $\mathbb{E}[\int_0^1 \frac{|W_1 - W_s|}{1-s}\, ds] < \infty$. Since the L^2–norm dominates the L^1–norm, we see that $\mathbb{E}[|W_1 - W_s|] \leq \mathbb{E}[(W_1 - W_s)^2]^{\frac{1}{2}} = \sqrt{1-s}$, and thus that $\mathbb{E}[\int_0^1 \frac{|W_1 - W_s|}{1-s}\, ds] \leq \int_0^1 \frac{1}{\sqrt{1-s}}\, ds < \infty$ as required.

Finally, if $t > 1$, then $\mathcal{F}_t = \mathcal{H}_t$. It is now easy to see that M is a $\mathbb{H}$–martingale.

$\square$

Example 1.2 Let W be a standard Brownian motion, with natural filtration $\mathbb{F}$, and define $\mathbb{G}$ by $\mathcal{G}_t = \mathcal{F}_{t+\varepsilon}$, where $\varepsilon > 0$. Then H' fails: W is not a $\mathbb{G}$–semimartingale.

To prove this, we use the Bichteler–Dellacherie Theorem, i.e. we show that W is not a good $(\mathbb{G}, \mathbb{P})$–integrator. It suffices to show that there exists a sequence H^n of $\mathbb{G}$–predictable elementary processes such that $H^n \to 0$ uniformly in probability on compacts, yet $(H^n \bullet W) \not\to 0$ ucp (cf. [**?**], Thm. II.11). Let $n_0 \in \mathbb{N}$ be least such that $2^{-n_0} \leq \varepsilon$. For $n \geq N$, define

$$H^n = \sum_{k=0}^{2^n-1} \Delta_k^n W I_{(\frac{k}{2^n}, \frac{k+1}{2^n}]} \qquad \text{where} \qquad \Delta_k^n W = W_{\frac{k+1}{2^n}} - W_{\frac{k}{2^n}}$$

Note that if $Y \sim N(0,1)$ and $y \geq 0$, then $\mathbb{P}(Y \geq y) \leq \frac{1}{\sqrt{2\pi}y} e^{-\frac{1}{2}y^2}$, a fact which follows easily from the identity $e^{-\frac{1}{2}y^2}\frac{1}{y} \leq \int_y^\infty e^{-\frac{1}{2}x^2}(1 + \frac{1}{x^2})\,dx$. It follows that, for $\delta > 0$, we have

$$\mathbb{P}(|\Delta_k^n W| > \delta) \leq \frac{2}{2^{n/2}\delta\sqrt{2\pi}}\, e^{-\frac{1}{2}2^n\delta^2}$$

and hence that

$$\mathbb{P}(\sup_{t\leq 1} |H_t^n| > \delta) \leq \sum_{k=0}^{2^n-1} \mathbb{P}(|\Delta_k^n W| > \delta) \leq \frac{2^n \cdot 2}{2^{n/2}\delta\sqrt{2\pi}}\, e^{-\frac{1}{2}2^n\delta^2} = \frac{KC}{e^{\frac{1}{2}K^2 D}}$$

where $K = 2^{n/2}, C = \frac{2}{\delta\sqrt{2\pi}}$ and $D = \delta^2$. It follows easily that $H^n \to 0$ ucp as $n \to \infty$. However,

$$(H^n \bullet W)_1 = \sum_{k=0}^{2^n-1} (\Delta_k^n W)^2$$

so that $\mathbb{E}(H^n \bullet W)_1 = 1$. Moreover, the family $\{(H^n \bullet W)_1 : n \geq n_0\}$ is clearly $\mathcal{L}^2$–bounded (by 5, for example), hence UI. We therefore cannot have $(H^n \bullet W) \to 0$ ucp.

□

Remarks 1.3 (1.) Protter[**?**], citing Ito, shows that Example 1.1 also applies to Lévy processes: If Z is a Lévy process on $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, and if $\mathcal{H}$ is the smallest filtration satisfying the usual conditions such that $\sigma(Z_1) \cup \mathcal{F}_t \subseteq \mathcal{H}_t$ for all $t \geq 0$, then Z is a $\mathbb{H}$–semimartingale. In addition, if $\mathbb{E}|Z_t| < \infty$ for all $t \geq 0$, then

$$M_t = Z_t - \int_0^{t\wedge 1} \frac{Z_1 - Z_s}{1-s}\, ds$$

is a $\mathbb{H}$–martingale.

(2.) Later, we will generalize Example 1.1 by considering enlargements of the form $\sigma(\int_0^\infty \varphi(s)\, dW_s)$.

□

Remarks 1.4 The semimartingale property is closely related to the no–arbitrage property. Delbaen and Schachermayer[**?**] show that if S is a locally bounded càdlàg asset price process satisfying the“no free lunch with vanishing risk property” (NFLVR) for simple portfolios (i.e. linear combinations of buy–and–hold portfolios $I_{(T_1,T_2]}$, where $T_1 \leq T_2$ are stopping times,

then S must be a semimartingale. In other words, if the semimartingale property is destroyed by enlargement of filtration, then there must be something very close to arbitrage for those with access to the enlarged information set.

However, the semimartingale property is not sufficient to guarantee no–arbitrage. As we have just seen, adding knowledge of W_T does not destroy the semimartingale property. But in the Black–Scholes model (with constant coefficients), knowing W_T is the same as knowing the terminal stock price S_T, and that clearly entails arbitrage.

□

Consider a probability space Let $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with right–continuous filtrations $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$. We are interested in the enlarged filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ which amalgamates the information in $\mathbb{F}, \mathbb{H}$. The natural candidate for this is $(\mathcal{F}_t \vee \mathcal{H}_t)_t := (\sigma(\mathcal{F}_t \cup \mathcal{H}_t)_t$. However, in order to do analysis, we need to work with càdlaàg versions of stochastic processes, and for those to exist, we require the usual hypotheses to hold. Hence we will define

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \mathcal{H}_s)$$

Two special cases of enlargement have received significant attention:

(1) If X is a random variable, and $\mathcal{H}_t = \sigma(X)$ for all t, then

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \sigma(X))$$

is an *initial enlargement* of $\mathbb{F}$ by X. The information $\sigma(X)$ is added to $\mathbb{F}$ all at once.

(2) If T is a random time, but not necessarily an $\mathbb{F}$–stopping time, then we can gradually add just enough information to make it a stopping time. Then

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \sigma(T \wedge s))$$

Here $\mathbb{G}$ is a *progressive enlargement* of $\mathbb{F}$.
(For example, an insider may not know exactly when an event (e.g. bankruptcy) is going to happen, but will know if it has happened. The time T of bankruptcy is then a stopping time for the insider, but not the "honest" investor.

In this chapter, we are mainly concerned with initial enlargements. nevertheless, we start off in a very general framework.

2 Enlargements of Filtrations via Changes of Measure

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with right–continuous filtrations $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$. The object of our studies is the enlarged filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$, where

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \mathcal{H}_s)$$

Following Ankirchner et al.([**?**],[**?**]), we find a "translation" of the current set–up in a product space, where it will transpire that the enlargement of filtration corresponds to a change of measure. Consider, therefore, the product space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, \bar{\mathbb{F}})$, where

$$\bar{\Omega} = \Omega \times \Omega \qquad \bar{\mathcal{F}} = \mathcal{F}_\infty \otimes \mathcal{H}_\infty \qquad \bar{\mathbb{P}} = \mathbb{P}i^{-1} \qquad \bar{\mathcal{F}}_t = \bigcap_{s>t} (\mathcal{F}_s \otimes \mathcal{H}_s)$$

and $i : \Omega \to \bar{\Omega} : \omega \mapsto (\omega, \omega)$ is the diagonal embedding. By the change of variable formula, it follows that

$$\int f(\omega, \omega')\, d\bar{\mathbb{P}}(\omega, \omega') = \int f(\omega, \omega)\, d\mathbb{P}(\omega)$$

for every $\bar{\mathcal{F}}$–measurable f.

Right now, our main concern is how various properties from stochastic analysis fare under the translation. We have two translation mappings, the diagonal embedding, and the first projection mapping

$$i : \Omega \to \bar{\Omega} : \omega \mapsto (\omega, \omega) \qquad \pi : \bar{\Omega} \to \Omega : (\omega, \omega') \mapsto \omega$$

The map i is used to translate from $\bar{\Omega}$ to Ω. We will show, roughly, that if $\bar{X}$ is measurable/predictable/semimartingale w.r.t $(\bar{\mathbb{F}}, \bar{\mathbb{P}})$ then $\bar{X} \circ i$ is measurable/predictable/semimartingale w.r.t $(\mathbb{G}, \mathbb{P})$. Conversely, the map π is used to translate from Ω to $\bar{\Omega}$, and we will see, roughly, that if X is a $(\mathbb{F}, \mathbb{P})$–semimartingale, then $X \circ \pi$ is a $(\bar{\mathbb{F}}, \bar{\mathbb{P}})$–semimartingale. We say "roughly", because we do have to keep track of null sets. We will be working with different measures, which therefore have different null sets. Now much of stochastic analysis works only when we impose the usual conditions on filtrations — e.g., these are required to prove that martingales have càdlàg versions — so that we must ensure that our filtrations are properly augmented w.r.t. the right measure, if we want to be able to do any analysis. Thus, given a filtration $\mathbb{K} = (\mathcal{K}_t)_{t \geq 0}$ and a probability measure P, we let $\mathbb{K}^P = (\mathcal{K}_t^P)_{t \geq 0}$ be the filtration $\mathbb{K}$ completed by P–negligible sets.

Lemma 2.1 *(a) We have*

$$\mathcal{G}_t = i^{-1}(\bar{\mathcal{F}}_t) \qquad \mathcal{G}_t^{\mathbb{P}} \supseteq i^{-1}(\bar{\mathcal{F}}_t^{\bar{\mathbb{P}}})$$

(b) If $\bar{X}_n \to 0$ in $\bar{\mathbb{P}}$–probability, then $\bar{X}_n \circ i \to 0$ in $\mathbb{P}$–probability.

(c) If $\bar{\tau}$ is a $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$–stopping time, then $\bar{\tau} \circ i$ is a $\mathbb{G}^{\mathbb{P}}$–stopping time.

Proof: (cf. [**?**]) (a) As $\mathcal{F}_t \vee \mathcal{H}_t = \sigma(F \cap H : F \in \mathcal{F}_t, H \in \mathcal{H}_t)$, we have

$$\begin{aligned}
\mathcal{G}_t &= \bigcap_{s>t} \sigma(F \cap H : F \in \mathcal{F}_s, H \in \mathcal{H}_s) \\
&= \bigcap_{s>t} \sigma(i^{-1}(F \times H) : F \in \mathcal{F}_s, H \in \mathcal{H}_s) \\
&= i^{-1}\left(\bigcap_{s>t} \mathcal{F}_s \otimes \mathcal{H}_s \right) \\
&= i^{-1}(\bar{\mathcal{F}}_t)
\end{aligned}$$

Now if $A \in \bar{\mathcal{F}}_t^{\bar{\mathbb{P}}}$, then there exists $B \in \bar{\mathcal{F}}_t$ such that $\bar{\mathbb{P}}(A \Delta B) = 0$. Hence $\mathbb{P}(i^{-1}(A) \Delta i^{-1}(B)) = 0$. Now $i^{-1}(B) \in \mathcal{G}_t$, and hence $i^{-1}(A) \in \mathcal{G}_t^{\mathbb{P}}$.

(b) Let $X_n = \bar{X}_n \circ i$. Then $\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(i^{-1}\{|\bar{X}_n| > \varepsilon\}) = \bar{\mathbb{P}}(|\bar{X}_n| > \varepsilon) \to 0$ as $n \to \infty$.
(c) If $\bar{\tau}$ is a $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$–stopping time, then

$$\{\bar{\tau} \circ i \leq t\} = i^{-1}\{\bar{\tau} \leq t\} \in i^{-1}(\bar{\mathcal{F}}_t^{\bar{\mathbb{P}}}) \subseteq \mathcal{G}_t^{\mathbb{P}}$$

so that $\bar{\tau} \circ i$ is a $\mathbb{G}^{\mathbb{P}}$–stopping time.

$\dashv$

The following result translates stochastic analysis from $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$ to $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$:

Theorem 2.2 *(a) If $\bar{\Omega} \xrightarrow{\bar{f}} \mathbb{R}$ is $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$–measurable, then $\bar{f} \circ i$ is $\mathbb{G}^{\mathbb{P}}$–measurable.*

(b) If $\bar{X}$ is $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$–adapted, then $\bar{X} \circ i$ is $\mathbb{G}^{\mathbb{P}}$–adapted.

(c) If $\bar{X}$ is $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$–predictable, then $\bar{X} \circ i$ is $\mathbb{G}^{\mathbb{P}}$–predictable.

(d) If $\bar{X}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–(local) martingale, then $\bar{X} \circ i$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–(local) martingale.

(e) If $\bar{X}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–semimartingale, then $\bar{X} \circ i$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–semimartingale.

(f) If $\bar{H}$ is càglàd $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–adapted and $\bar{X}$ a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–semimartingale, and $H = \bar{H} \circ i, X = \bar{X} \circ i$, then

$$(\bar{H} \bullet \bar{X}) \circ i = H \bullet X$$

up to indistinguishability.

(g) If $\bar{X}, \bar{Y}$ are $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–semimartingales, and $X = \bar{X} \circ i, Y = \bar{Y} \circ i$, then

$$[\bar{X}, \bar{Y}] \circ i = [X, Y]$$

up to indistinguishability.

Proof: (a) follows from Lemma 2.1 by the usual arguments, and (b) follows immediately from (a).
(c) Consider first an $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–predictable process of the form $\bar{X} = I_A I_{(s,t]}$, where $A \in \bar{\mathcal{F}}_s^{\bar{\mathbb{P}}}$. Then $I_A \circ i$ is $\mathcal{G}_s^{\mathbb{P}}$–measurable, by (a), and so $\bar{X} \circ i = (I_A \circ i) I_{(s,t]}$ is $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–predictable. The result for general predictable $\bar{X}$ now follows using a monotone class theorem.
(d) Suppose first that $\bar{X}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–martingale. Put $X = \bar{X} \circ i$ and let $0 \leq s < t, G \in \mathcal{G}_s^{\mathbb{P}}$. To show that X is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–martingale, it suffices to show that $\mathbb{E}_{\mathbb{P}}[I_G(X_t - X_s)] = 0$. Now by Lemma 2.1, there is $\bar{F} \in \bar{\mathcal{F}}_s^{\bar{\mathbb{P}}}$ such that $G = i^{-1}(\bar{F})$. Then

$$\mathbb{E}_{\mathbb{P}}[I_G(X_s - X_t)] = \int (I_{\bar{F}}(\bar{X}_t - \bar{X}_s)) \circ i \; d\mathbb{P} = \int I_{\bar{F}}(\bar{X}_t - \bar{X}_s) \; d\bar{\mathbb{P}} = \mathbb{E}_{\bar{\mathbb{P}}}[I_{\bar{F}}(\bar{X}_t - \bar{X}_s)] = 0$$

If $\bar{X}$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–local martingale, the result follows easily by localization, using (d).
(e) We sketch the proof. Suppose that $\bar{X}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–semimartingale, and put $X = \bar{X} \circ i$. Clearly, X is càdlàg, because $\bar{X}$ is, and $\mathbb{G}^{\mathbb{P}}$–adapted. To prove that X is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–semimartingale, it suffices to show that if $(\theta^n)_n$ is a sequence of simple predictable processes converging to zero uniformly in (t, ω), then $(\theta^n \bullet Y)$ converges to zero in $\mathbb{P}$–probability, by the Bichteler–Dellacherie Theorem. Now if $\theta = \sum_{i=1}^n \theta_i I_{(t_i, t_{i+1}]}$ is simple, where each θ_i is $\mathcal{G}_{t_i}$–measurable, then $\theta_i = \bar{\theta}_i \circ i$ for some $\bar{\mathcal{F}}_{t_i}$–measurable $\bar{\theta}_i$. Consequently, for every simple

$\mathbb{G}$–predictable sequence $(\theta^n)_n$ converging to 0 uniformly, we can find a simple $\bar{\mathbb{F}}$–predictable sequence $(\bar{\theta}^n)_n$ converging to 0 uniformly such that $\theta^n = \bar{\theta}^n \circ i$: Since $\bar{X}$ is a semimartingale, $(\bar{\theta}^n \bullet \bar{X})$ converges to 0 in $\bar{\mathbb{P}}$–probability, and hence $(\theta^n \bullet X)$ converges to 0 in $\mathbb{P}$–probability.
(f) For $n \in \mathbb{N}$, and $1 \leq i \leq 2^n$, let $t_i^n = ti2^{-n}$. Then

$$\bar{H}_0 \bar{X}_0 + \sum_{i=0}^{2^n-1} \bar{H}_{t_i^n}(\bar{X}_{t_{i+1}^n} - \bar{X}_{t_i^n})$$

converges to $(\bar{H} \bullet \bar{X})_t$ in $\bar{\mathbb{P}}$–probability, as $n \to \infty$. Thus $(\bar{H} \bullet \bar{X})_t \circ\ i$ is the limit in $\mathbb{P}$–probability of

$$H_0 X_0 + \sum_{i=0}^{2^n-1} H_{t_i^n}(X_{t_{i+1}^n} - X_{t_i^n})$$

This limit is also $(H \bullet X)_t$, and so $(\bar{H} \bullet \bar{X}) \circ\ i$ and $(H \bullet X)$ are modifications of each other. Since both are càdlàg, they must be indistinguishable under $\mathbb{P}$.
(g) follows in a similar way, using the fact that $[\bar{X}, \bar{Y}]_t$ is a limit in probability of $\bar{X}_0 \bar{Y}_0 + \sum_{i=0}^{2^n-1}(\bar{X}_{t_{i+1}^n} - \bar{X}_{t_i^n})(\bar{Y}_{t_{i+1}^n} - \bar{Y}_{t_i^n})$

$\dashv$

To translate stochastic analysis from Ω to $\bar{\Omega}$, we don't need quite as much. Given a probability measure R on $(\Omega, \mathcal{F})$, let

$$\bar{\mathbb{Q}} = \mathbb{P}|_{\mathcal{F}_\infty} \otimes R|_{\mathcal{H}_\infty}$$

$\bar{\mathbb{Q}}$ is called a *decoupling* measure. We shall often take $R = \mathbb{P}$.

Lemma 2.3 *If $A \in \mathcal{F}_t^{\mathbb{P}}$, then $A \times \Omega \in \bar{\mathcal{F}}_t^{\bar{\mathbb{Q}}}$.*

Proof: Take B in $\mathcal{F}_t$ so that $\mathbb{P}(A \Delta B) = 0$, and note that $\bar{\mathbb{Q}}((A \times \Omega)\Delta(B \times \Omega)) = \mathbb{P}(A \Delta B) = 0$.

$\dashv$

Theorem 2.4 *Let M be a right–continuous (everywhere) $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$–(local) martingale. Then $M \circ \pi$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–(local) martingale.*

Proof: Let $\bar{M} = M \circ \pi$, i.e. $\bar{M}(\omega, \omega') = M(\omega)$. It is clear that $\bar{M}$ is right–continuous, because M is. To see that $\bar{M}$ is $\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}$–adapted, note that $c \in \mathbb{R}$, we have

$$\{\bar{M}_t \leq c\} = \{M_t \leq c\} \times \Omega \in \bar{\mathcal{F}}_t^{\bar{\mathbb{Q}}}$$

by the lemma, which proves that $\bar{M}$ is adapted.

Suppose first that M is a martingale. For $0 \leq s < t$, $A \in \mathcal{F}_s, B \in \mathcal{H}_s$, we have

$$\mathbb{E}_{\bar{\mathbb{Q}}}[I_{A \times B}(\bar{M}_t - \bar{M}_s)] = R(B)\mathbb{E}_{\mathbb{P}}[I_A(M_t - M_s)] = 0$$

By a monotone class argument, it follows that $\mathbb{E}_{\bar{\mathbb{Q}}}[\bar{X}(\bar{M}_t - \bar{M}_s)] = 0$ for all bounded $\mathcal{F}_s \otimes \mathcal{H}_s$–measurable $\bar{X}$. As $\bar{M}$ is right–continuous, we see that $\mathbb{E}_{\bar{\mathbb{Q}}}[\bar{X}(\bar{M}_t - \bar{M}_s)] = 0$ for all bounded $\bigcap_{u>s}(\mathcal{F}_u \otimes \mathcal{H}_u)$–measurable $\bar{X}$, which proves that $\bar{M}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–martingale.

The result can now easily be extended (via localization) to the case where M is a local martingale: If τ is a $\mathbb{F}^{\mathbb{P}}$–stopping time, and $\bar{\tau} = \tau \circ \pi$, then

$$\{\bar{\tau} \leq t\} = \{\tau \leq t\} \times \Omega \in \bar{\mathcal{F}}_t^{\bar{\mathbb{Q}}}$$

by the lemma, so that $\bar{\tau}$ is a $\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}$–stopping time.

⊣

We shall henceforth impose the following assumption on $\bar{\mathbb{P}}, \bar{\mathbb{Q}}$:

Assumption A:	$\bar{\mathbb{P}} << \bar{\mathbb{Q}}$	on	$\bar{\mathcal{F}}$

Now recall that the semimartingale property is preserved under a change of measure (provided the new measure is absolutely continuous w.r.t. the original). In particular, every $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–semimartingale is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{P}})$–semimartingale. We therefore immediately obtain the following important result:

Theorem 2.5 (**A**) *implies* (**H′**).
I.e. if $\bar{\mathbb{P}} << \bar{\mathbb{Q}}$*, then every* $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$*–semimartingale is a* $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$*–semimartingale.*

Proof: Suppose that M is a $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$–local martingale. It suffices to show that it is also a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–semimartingale. Now $\bar{M} = M \circ \pi$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–local martingale, by Theorem 2.4. Since $\bar{\mathbb{P}} << \bar{\mathbb{Q}}$, we have that $\bar{M}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{P}})$–semimartingale. By Stricker's Theorem, it is also a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–semimartingale (because $\bar{M}$ is adapted to $\bar{\mathbb{F}}^{\bar{\mathbb{P}}} \subseteq \bar{\mathbb{F}}^{\bar{\mathbb{Q}}}$). Hence $\bar{M} \circ i = M \circ \pi \circ i = M$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–semimartingale, by Theorem 2.2.

⊣

3 Doob–Meyer Decompositions via Girsanov's Theorem

We have now accomplished the following: Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtrations $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ satisfying the usual conditions, we have shown that assumption (**A**) implies hypothesis (**H′**) for the enlarged filtration $\mathbb{G} = (\bigcap_{s>t}(\mathcal{F}_s \vee \mathcal{H}_s))_{t\geq 0}$. In essence, ignoring negligible sets, we embedded the original set–up $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ into a product space

$$(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{Q}}, \bar{\mathbb{F}}) = (\Omega \times \Omega, \mathcal{F}_\infty \otimes \mathcal{H}_\infty, \mathbb{P}|\mathcal{F}_\infty \otimes \mathbb{P}|\mathcal{H}_\infty, (\bigcap_{s>t} \mathcal{F}_s \otimes \mathcal{H}_s)_{t\geq 0})$$

The first projection $\pi : (\bar{\Omega}, \bar{\mathcal{F}}_t, \bar{\mathbb{Q}}) \to (\Omega, \mathcal{F}_t, \mathbb{P}) : (\omega, \omega') \mapsto \omega$ is clearly measurable (for each $t \geq 0$), i.e. the first component captures the initial set–up $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. Theorem 2.2 however, shows that the enlarged set–up $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G})$ is captured perfectly by the space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}, \bar{\mathbb{F}})$. In essence, therefore, enlarging filtrations on Ω corresponds to changing measures on $\bar{\Omega}$. This means that we can bring the machinery of Girsanov's Theorem into play. Again, we rely entirely on [**?**],[**?**].

Since we are assuming that $\bar{\mathbb{P}} << \bar{\mathbb{Q}}$ on $\bar{\mathcal{F}}$, let $\bar{Z}$ be a càdlàg version of the likelihood process:

$$\bar{Z}_t = \frac{d\bar{\mathbb{P}}}{d\bar{\mathbb{Q}}}\Big|\bar{\mathcal{F}}_t^{\bar{\mathbb{Q}}}$$

(The regularization theorems guaranteeing the existence of a càdlàg version require a right–continuous and complete filtration.) As $\bar{Z}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–martingale, it follows that the process $Z = \bar{Z} \circ i$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–semimartingale.

Now the Lenglart–Girsanov theorem states that every $\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–semimartingale is also a $\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{P}})$–semimartingale. To be precise, let $\bar{M}$ be a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–local martingale with $\bar{M}_0 = 0$. Define

$$\bar{T} = \inf\{t > 0 : \bar{Z}_t = 0, \bar{Z}_{t-} > 0\} \qquad \bar{U}_t = \Delta \bar{M}_{\bar{T}} I_{\{t \geq \bar{T}\}}$$

and let $\tilde{U}$ be the $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–compensator of $\bar{U}$ (i.e. the *unique predictable* finite variation process such that $\bar{U} - \tilde{U}$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–local martingale; cf. Protter[**?**], III.5). By the Lenglart–Girsanov Theorem (cf. Protter[**?**], III.8), the process

$$\bar{M}_t - \int_0^t \frac{1}{\bar{Z}_s}\, d[\bar{M}, \bar{Z}]_s + \tilde{U}_s$$

is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–local martingale. Now let $U = \tilde{U} \circ i$, and recall that $Z = \bar{Z} \circ i$. Then:

Theorem 3.1 *If M is a $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$–local martingale with $M_0 = 0$, then*

$$M - \frac{1}{Z} \bullet [Z, M] + U$$

is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–local martingale.

Proof: By Theorem 2.4, the process $\bar{M} = M \circ \pi$ is a $(\bar{\mathbb{F}}^{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}})$–local martingale with $\bar{M}_0 = 0$. By the Lenglart–Girsanov Theorem,

$$\bar{X} = \bar{M} - \frac{1}{\bar{Z}} \bullet [\bar{Z}, \bar{M}] + \tilde{U}$$

is a $(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$–local martingale. By Theorem 2.2, $\bar{X} \circ i$ is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–local martingale. But, by the same theorem, we see that

$$X \circ i = M - \frac{1}{Z} \bullet [Z, M] + U$$

$\dashv$

If M is a continuous $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$–local martingale with $M_0 = 0$, and $\bar{M} = M \circ \pi$, etc., then clearly $U = 0$ (because $\bar{U} = 0$, and hence $\tilde{U} = 0$), so that we have:

Theorem 3.2 *If M is a continuous $(\mathbb{F}^{\mathbb{P}}, \mathbb{P})$–local martingale with $M_0 = 0$, then*

$$M - \frac{1}{Z} \bullet [Z, M]$$

is a $(\mathbb{G}^{\mathbb{P}}, \mathbb{P})$–local martingale.

4 Jacod's Criterion for Initial Enlargements

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a filtered probability space satisfying the usual hypotheses, and let X be an $\mathcal{F}$–measurable random element with values in a state space $(S, \mathcal{S})$. We consider an initial enlargement $\mathbb{G}$ of $\mathbb{F}$, given by

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \sigma(X))$$

i.e.

$$\mathcal{G}_t = \bigcap_{s>t} (\mathcal{F}_s \vee \mathcal{H}_s) \qquad \text{where} \qquad \mathcal{H}_s = \sigma(X) \qquad \text{for all } s \geq 0$$

The most well–known condition for an initial enlargement to satisfy condition $(\mathbf{H}')$ is *Jacod's Criterion* (due to Jacod[**?**]), which we shall now prove.

First recall that if X is a random element with values in a Borel space $(S, \mathcal{S})$, then X has a (unique) regular conditional distribution $Q_t(\omega, dx)$ with the properties that

(i) $A \mapsto Q_t(\omega, A)$ is a probability measure on $(S, \mathcal{S})$, for almost all ω.

(ii) $\omega \mapsto Q_t(\omega, A)$ is measurable for all $A \in \mathcal{S}$;

(iii) $Q_t(\cdot, dx)$ is a version of $\mathbb{P}(X \in dx|\mathcal{F}_t)$.

Jacod's criterion depends on the following assumption:

> **Assumption J:** For each $t \geq 0$, here exists a σ–finite measure $\eta_t(ds)$ on the state space $(S, \mathcal{S})$ of X such that
>
> $$Q_t(\omega, ds) << \eta_t(ds) \quad \mathbb{P}\text{–a.s.}$$

Theorem 4.1 (Jacod's Criterion*) Let X be a random element on $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with regular conditional distribution $Q_t(\omega, dx)$. If* $(\mathbf{J})$ *holds, then* $(\mathbf{H}')$ *holds for the initial enlargement of* $\mathbb{F}$ *with* $\sigma(X)$.

We prove this result in two steps. First we show that $(\mathbf{J})$ is equivalent to a seemingly stronger statement $(\mathbf{J}')$:

> **Assumption J′:** there exists a *single* σ–finite measure $\eta(ds)$ on the state space $(S, \mathcal{S})$ of X such that
>
> $$Q_t(\omega, ds) << \eta(ds) \qquad \text{for all } \omega \in \Omega \text{ and } t \geq 0$$

Then we show that $(\mathbf{J}')$ is equivalent to $(\mathbf{A})$.

Lemma 4.2 *In the set–up of Theorem 4.1, if* $(\mathbf{J})$ *holds, then so does* $(\mathbf{J}')$.
In fact we may take the single *σ–finite measure η on the state space $(S, \mathcal{S})$ of X satisfying*

$$Q_t(\omega, dx) << \eta(dx) \qquad \text{for all } \omega \in \Omega \text{ and } t \geq 0$$

to be the law of X.

Proof: We follow Protter[?]. Assume that $(\mathbf{J}')$ holds,, and that $Q_t(\omega, ds) << \eta_t(ds)$ for all ω. By Theorem C.8 there exists a $\mathcal{F}_t \otimes \mathcal{S}$–measurable function $q_t(\omega, s)$ such that $Q_t(\omega, ds) = q_t(\omega, s)\, \eta(ds)$. Define $a_t(s) = \mathbb{E}[q_t(\cdot, s)]$ and define

$$r_t(\omega, s) = \begin{cases} \dfrac{q_t(\omega, s)}{a_t(s)} & \text{if } a_t(s) > 0 \\ 0 & \text{else} \end{cases}$$

Since we must have $q_t(\omega, s) = 0$ a.s. whenever $a_t(s) = 0$, it follows that $q_t(\omega, s) = r_t(\omega, s)a_t(s)$ a.s. Hence

$$Q_t(\omega, ds) = r_t(\omega, s)a_t(s)\eta_t(ds) \qquad \text{a.s.}$$

Now let η be the law of X. For every non–negative $\mathcal{S}$–measurable function g, and every $t \geq 0$, we have

$$\begin{aligned}\int g(s)\ \eta(ds) &= \mathbb{E}[g(X)] \\ &= \mathbb{E}\left[\int g(s)\ Q_t(\cdot, ds)\right] \\ &= \mathbb{E}\left[\int g(s) q_t(\cdot, s)\eta_t(ds)\right] \\ &= \int g(s)\mathbb{E}[q_t(\cdot, s)]\ \eta_t(ds) \\ &= \int g(s) a_t(s)\ \eta_t(ds)\end{aligned}$$

Hence

$$a_t(s)\eta_t(ds) = \eta(ds) \qquad \text{for all } t \geq 0$$

so that $Q_t(\omega, ds) = r_t(\omega, s)\eta(ds)$, as required.

$\dashv$

Lemma 4.3 *(i) The map $j : (\Omega \times \Omega, \mathcal{F}_t \otimes \sigma(X)) \to (\Omega \times S, \mathcal{F}_t \otimes \mathcal{S}) : (\omega, \omega') \mapsto (\omega, X(\omega'))$ is measurable.*

(ii) Every $C \in \mathcal{F}_t \otimes \sigma(X)$ is of the form $j^{-1}(D)$ for some $D \in \mathcal{F}_t \otimes \mathcal{S}$.

Proof: (i) follows from the fact that $j^{-1}(A \times B) = A \times X^{-1}(B)$.
(ii) The family of all C which can be represented as $j^{-1}(D)$, is a σ–algebra containing all the measurable rectangles in $\mathcal{F}_t \otimes \sigma(X)$.

$\dashv$

Lemma 4.4 *If $f : (\Omega \times S, \mathcal{F}_t \otimes \mathcal{S}) \to \mathbb{R}$ is measurable, then*

$$\int_{\Omega\times\Omega} f(\omega, X(\omega'))\ \bar{\mathbb{P}}(d\omega) = \int_\Omega \int_S f(\omega, s) Q_t(\omega, ds)\ \mathbb{P}(d\omega)$$

Proof: Define a probability measure $\tilde{P}$ on $\mathcal{F}_t \otimes \sigma(X)$ by

$$\tilde{P}(A \times X^{-1}(B)) = \int_A Q_t(\omega, B)\ \mathbb{P}(d\omega) \qquad \text{for } A \in \mathcal{F}_t, B \in \mathcal{S}$$

$\dashv$

On such rectangles $A \times X^{-1}(B)$, we have

$$\tilde{P}(A \times X^{-1}(B)) = \mathbb{E}[Q_t(\cdot, B) I_A] = \mathbb{E}[I_{X\in B} I_A] = \mathbb{P}(A \cap X^{-1}(B)) = \bar{\mathbb{P}}(A \times X^{-1}(B))$$

Consequently, $\tilde{P} = \bar{\mathbb{P}}$.

It is now easy to verify that

$$\int f(\omega, X(\omega'))\ \tilde{P}(d\omega \times d\omega') = \iint f(\omega, s)\ Q_t(\omega, ds)\ \mathbb{P}(d\omega)$$

This follows directly from the definition of $\tilde{P}$ if $f = I_{A\times B}$ is the indicator of a measurable rectangle, and the follows for arbitrary measurable f by the usual arguments.

$\dashv$

Proposition 4.5 *In the set–up of Theorem 4.1,* $(\mathbf{J}')$ *is equivalent to* $(\mathbf{A})$*: To be precise, let* R *be a probability measure with* $RX^{-1} = \eta$*, where* η *is the emasure supplied by* $(\mathbf{J}')$*. (In particular, we may take* $R = \mathbb{P}$*, because we may take* η *to be* $\mathbb{P}X^{-1}$*, by Lemma 4.2.)*
With $\mathcal{H}_t = \sigma(X)$*, we have*

$$\bar{\mathbb{P}} << \bar{\mathbb{Q}} := \mathbb{P} \otimes \mathbb{R} \quad \textit{on } \bar{\mathcal{F}}_t \textit{ for all } t \geq 0$$

if and only if

$$Q_t(\omega, ds) << \eta(ds) \quad \textit{for almost all } \omega\textit{, for all } t \geq 0$$

Proof: ([?], [?]) First assume $(\mathbf{J}')$, and thus that the regular conditional version $Q_t(\omega, ds)$ of $\mathbb{P}(X \in ds|\mathcal{F}_t)$ is absolutely continuous w.r.t. the law η of X, for each $t \geq 0$ — cf. Lemma 4.2. Let $t \geq 0$.

Now, for $s < t$, let $C \in \bar{\mathcal{F}}_s = \bigcap_{u>s}(\mathcal{F}_u \otimes \sigma(X))$ with $\bar{\mathbb{Q}}(C) = 0$. By Lemma 4.3, there is $D \in \mathcal{F}_t \otimes \mathcal{S}$ such that $C = j^{-1}(D)$. Then, since $\bar{\mathbb{Q}} = \mathbb{P} \otimes \mathbb{P}$ (restricted to suitable σ–algebras)

$$\begin{aligned}\bar{\mathbb{Q}}(C) &= \int I_D \circ j(\omega, \omega')\ \bar{\mathbb{Q}}(d\omega \times d\omega') \\ &= \int_\Omega \left(\int_S I_D(\omega, s)\ \eta(ds) \right)\ \mathbb{P}(d\omega) = 0\end{aligned}$$

The inner integral (being non–negative) must therefore satisfy $\int_S I_D(\omega, s)\ \eta(ds) = 0$ for $\mathbb{P}$–a.a. ω. As $Q_t(\omega, ds) << \eta(ds)$, also $\int_S I_D(\omega, s)\ Q_t(\omega, ds) = 0$ for almost all t, ω. Hence by Lemma 4.4,

$$\bar{\mathbb{P}}(C) = \int_\Omega I_D(\omega, X(\omega'))\ \bar{\mathbb{P}}(d\omega \times d\omega') = \int_\Omega \int_S I_D(\omega, s) Q_t(\omega, ds)\ \mathbb{P}(d\omega) = 0$$

This proves that $(\mathbf{J}') \Rightarrow (\mathbf{A})$.

Conversely, suppose that $\bar{\mathbb{P}} << \bar{\mathbb{Q}} = \mathbb{P} \otimes R$ on $\bar{\mathcal{F}}_t$. If $\varphi(\omega, \omega')$ is a $\mathcal{F}_t \otimes \sigma(X)$–measurable version of $\frac{d\bar{\mathbb{P}}}{d\bar{\mathbb{Q}}}$, then by lemma 4.3, there is a $\mathcal{F}_t \otimes \mathcal{S}$–measurable $\tilde{\varphi}$ such that $\varphi(\omega, \omega') = \tilde{\varphi}(\omega, X(\omega'))$. Define $U_t(\omega, B) = \int_{X^{-1}(B)} \tilde{\varphi}(\omega, X(\omega'))\ R(d\omega')$. It is easy to see that U_t is a stochastic kernel from $(\Omega, \mathcal{F}_t)$ to $(S, \mathcal{S})$. We claim that U_t is a regular conditional distribution of X, i.e. a version of $\mathbb{P}(X \in B|\mathcal{F}_t) = \mathbb{E}[I_{X^{-1}(B)}|\mathcal{F}_t]$. For if $A \in \mathcal{F}_t$, then

$$\begin{aligned}&\int_A \int_{X^{-1}(B)} \tilde{\varphi}(\omega, X(\omega'))\ R(d\omega')\mathbb{P}(d\omega) \\ &= \int I_{A \times X^{-1}(B)} \varphi\ d\bar{\mathbb{Q}} \\ &= \bar{\mathbb{P}}(A \times X^{-1}(B)) \\ &= \mathbb{P}(A \cap X^{-1}(B)) \\ &= \int_A I_{X^{-1}(B)}\ d\mathbb{P}\end{aligned}$$

Hence $U_t(\omega, ds) = Q_t(\omega, ds)$ $\mathbb{P}$–a.s. Now if $\eta(B) = 0$, then

$$U_t(\omega, B) = \int I_B(X(\omega') \tilde{\varphi}(\omega, X(\omega'))\ R(d\omega) = \int_B \tilde{\varphi}(\omega, s)\ \eta(ds) = 0$$

so that $Q_t(\omega, ds) << \eta(ds)$.

⊣

Proof of Theorem 4.1: Combine Theorem 2.5 and Proposition 4.5.

⊣

Remarks 4.6 (a) Alternate proofs of Jacod's criterion for ($\mathbf{H}'$) may be found in Jacod[**?**] and Protter[**?**].

(b) Though much is made of Jacod's criterion in the literature, it fails already in the simplest case, when we enlarge the natural filtration of Brownian motion W by W_T. Indeed, ($\mathbf{H}'$) fails in this case. Cf. Remarks 6.5.

□

We work in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ satisfying the usual conditions.

Corollary 4.7 *Let X be a random element (with state space a Borel space $(S, \mathcal{S})$), such that X is independent of $\mathbb{F}$, and let $\mathbb{G}$ be the enlargement of $\mathbb{F}$ by X. Then* ($\mathbf{J}'$) *holds.*
Hence every $\mathbb{F}$–semimartingale is a $\mathbb{G}$–semimartingale.

Proof: By independence, the regular conditional distributions $Q_t(\omega, ds)$ of X are equal to the law $\eta(ds)$ of X.

⊣

Corollary 4.8 *Let X be a discrete random variable (i.e. takes only countably many values), and let $\mathbb{G}$ be the enlargement of $\mathbb{F}$ by X. Then* ($\mathbf{A}$) *holds.*
Hence every $\mathbb{F}$–semimartingale is a $\mathbb{G}$–semimartingale.

Proof: If X has range $\{x_n : n \in \mathbb{N}\}$, then $\bar{\mathcal{F}} = \mathcal{F}_\infty \otimes \sigma(X)$ consists of countable unions of rectangles of the form $F \times \{X = x_n\}$, where $F \in \mathcal{F}_\infty$. With $\bar{\mathbb{Q}} = \bar{\mathbb{P}}|_{\mathcal{F}_\infty} \otimes \bar{\mathbb{P}}|_{\sigma(X)}$, we see that $\bar{\mathbb{Q}}(F \times \{X = x_n\}) = 0 \Rightarrow \mathbb{P}(F)\mathbb{P}(X = x_n) = 0 \Rightarrow \mathbb{P}(F \cap \{X = x_n\}) = \bar{\mathbb{P}}(F \times \{X = x_n\}) = 0$.

⊣

Corollary 4.9 (Jacod's Countable Enlargement) *Let $\mathcal{A} = \{A_n : n \in \mathbb{N}\}$ be a family of mutually disjoint events in $\mathcal{F}$, and let $\mathbb{G}$ be the enlargement of $\mathbb{F}$ by $\sigma(\mathcal{A})$.*
Then every $\mathbb{F}$–semimartingale is a $\mathbb{G}$–semimartingale.

Proof: Apply Corollary 4.8 to $X = \sum_n nI_{A_n}$.

⊣

5 Stochastic Integrals under Enlargements

This short section is devoted to a result of Jeulin[?] which gives necessary and sufficient conditions for the stochastic integral $H \bullet M$ of a $(\mathbb{F}, \mathbb{P})$–local martingale w.r.t a $\mathbb{F}$–predictable process H to be a $\mathbb{G}$–semimartingale. We rely on the exposition in Protter[?].

Recall that a process X is said to be locally integrable w.r.t $(\mathbb{F}, \mathbb{P})$ iff there exists an increasing sequence of $\mathbb{F}$–stopping times $T_n \uparrow \infty$ a.s. such that $\mathbb{E}[|X_{T_n}|; T_n > 0] < \infty$ for all $n \in \mathbb{N}$. Recall further that an $(\mathbb{F}, \mathbb{P})$–semimartingale X is special iff the process $X_t^* = \sup_{s \leq t} |X_s|$ is locally integrable. In particular, since any local martingale is obviously special, we have that M^* is locally integrable for any local martingale M. If $\mathbb{G}$ is an enlargement of $\mathbb{F}$, then every $\mathbb{F}$–stopping time is a $\mathbb{G}$–stopping time, and hence any $(\mathbb{F}, \mathbb{P})$–locally integrable process is $(\mathbb{G}, \mathbb{P})$–locally integrable. it therefore follows that:

Proposition 5.1 *If an $(\mathbb{F}, \mathbb{P})$–local martingale is a $(\mathbb{G}, \mathbb{P})$–semimartingale, where $\mathbb{G}$ is an enlargement of $\mathbb{F}$, then it is a special $(\mathbb{G}, \mathbb{P})$–semimartingale.*

□

Theorem 5.2 *Let M be an $(\mathbb{F}, \mathbb{P})$–local martingale, and let H be $\mathbb{F}$–predictable such that $(\int_0^t H_s^2 \; d[M]_s)_{t \geq 0}$ is locally integrable. Suppose that $\mathbb{G}$ is an enlargement of $\mathbb{F}$ such that M remains a $(\mathbb{G}, \mathbb{P})$–semimartingale. Then M is a special $(\mathbb{G}, \mathbb{P})$–semimartingale. If $M = \tilde{M} + A$ is its $(\mathbb{G}, \mathbb{P})$–canonical decomposition, then $H \bullet M$ is a $(\mathbb{G}, \mathbb{P})$–semimartingale iff $(\int_0^t H_s \, dA_s)_{t \geq 0}$ exists as a path–by–path Lebesgue–Stieltjes integral.*
In that case, the $(\mathbb{G}, \mathbb{P})$–canonical decomposition of $H \bullet M$ is $H \bullet M = H \bullet \tilde{M} + H \bullet A$.

Proof: LATER... Cf. Protter[?] or Jeulin[?]

⊣

6 Initial Enlargements in the Brownian World

We follow Yor[?], Mansuy and Yor[?]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space that supports a Brownian motion W, with natural filtration $\mathbb{F}$. Let X be an $\mathcal{F}_\infty$–measurable random variable, and let $\mathbb{G}$ be the enlargement of $\mathbb{F}$ by X. Given a bounded Borel function $f : \mathbb{R} \to \mathbb{R}$, let $\lambda(f) = \lambda_t(f))_{t \geq 0}$ be a continuous version of the martingale $(\mathbb{E}[f(X)|\mathcal{F}_t])_{t \geq 0}$. By the martingale representation theorem, there exists a unique predictable process $\dot{\lambda}(f)$ such that

$$\lambda_t(f) = \mathbb{E}[f(X)] + \int_0^t \dot{\lambda}_u(f) \; dW_u$$

Denote by $\lambda_t(dx)$ the regular conditional distributions of X w.r.t. $\mathcal{F}_t$, i.e. $\lambda_t(A)$ is a version of $\mathbb{P}(X \in A|\mathcal{F}_t)$. Then

$$\lambda_t(f) = \int f(x) \; \lambda_t(dx)$$

We now make the following assumption:

Assumption: There is a family $\dot{\lambda}_t(dx)$ of measures such that

$$\dot{\lambda}_t(f) = \int f(x) \; \dot{\lambda}_t(dx) \qquad t\text{–a.e.}$$

Theorem 6.1 *Assume that $\dot\lambda_t(dx) << \lambda_t(dx)\ dt \times d\mathbb{P}$–a.e., and define $\rho(x,s)$ by $\dot\lambda_t(dx) = \rho(x,t)\ \lambda_t(dx)$ Then for any $(\mathbb{F},\mathbb{P})$–martingale $M = \int_0^\cdot m_u\ dW_u$ there exists a $(\mathbb{G},\mathbb{P})$–local martingale $\tilde M$ such that*

$$M = \tilde M + \int_0^\cdot \rho(X,u)\ d[M,W]_u$$

provided that

$$\int_0^t |\rho(X,u)|\ |d[M,W]_u| < \infty \quad a.s.\ for\ t \geq 0$$

Proof: By the martingale representation theorem there is a predictable m such that $M = \int_0^\cdot m_u\ dW_u$. Let f be a bounded Borel function, and let $s < t$ and $F \in \mathcal{F}_s$. Then

$$\begin{aligned}
&\mathbb{E}[I_F f(X)(M_t - M_s)] \\
&= \mathbb{E}[I_F(\lambda_t(f)M_t - \lambda_s(f)M_s)] \\
&= \mathbb{E}\left[\int_s^t m_u \dot\lambda_u(f)\ du\right] \\
&= \mathbb{E}\left[I_F \int_s^t m_u \left(\int \rho(x,u) f(x)\ \lambda_u(dx)\right)\ du\right] \\
&= \mathbb{E}\left[I_F \int_s^t m_u f(X)\rho(X,u)\ du\right] \\
&= \mathbb{E}\left[I_F f(X) \int_s^t \rho(X,u)\ d[M,W]_u\right]
\end{aligned}$$

because $d[M,W]{=}m_t\ dt$. Hence $\mathbb{E}[I_F f(X)(M_t - M_s)] = \mathbb{E}\left[I_F f(X) \int_s^t \rho(X,u)\ d[M,W]_u\right]$ for all $F \in \mathcal{F}_s$ and every bounded Borel function f. By a monotone class argument, we see that

$$\mathbb{E}\left[M_t - M_s - \int_s^t \rho(X,u)\ d[M,W]_u\ \middle|\mathcal{G}_s\right] = 0$$

$\dashv$

Applying this result to $M = W$, and using Lévy's characterization of Brownian motion, we get immediately.

Corollary 6.2 *If, in Theorem 6.1, the $\mathbb{F}$–Brownian motion W decomposes as*

$$W = \tilde W + \int_0^\cdot \rho(X,s)\ ds$$

where $\tilde W$ is a $\mathbb{G}$–Brownian motion (provided that $\int_0^t |\rho(X,s)|\ ds < \infty$ a.s. for all $t \geq 0$).

The following result will allow for explicit computations:

Corollary 6.3 *Assume that $\lambda_t(dx) = \phi(t,x)\ dx$, where $\phi(t,x)$ has the form*

$$\phi(t,x) = \phi(0,x) \exp\left(\int_0^t \rho(x,s)\ dW_s - \frac{1}{2}\int_0^t \rho(x,s)^2\ ds\right)$$

Then $\dot\lambda_t(dx) = \rho(x,t)\ \lambda_t(dx)$, so that we may apply Theorem 6.1.

Proof: Since $\lambda_t(f) = \int f(x)\phi(t,x)\, dx$, we see that $d\lambda_t(f) = \int f(x)(\phi(t,x)\rho(x,t)\, dW_t)\, dx$. Hence

$$\lambda_t(f) = \lambda_0(f) + \int_0^t \int f(x)\phi(u,x)\rho(x,u)\, dx\, dW_u$$

so that $\dot{\lambda}_u(f) = \int f(x)\phi(u,x)\rho(x,u)\, dx$, and thus $\dot{\lambda}_t(dx) = \phi(t,x)\rho(x,t)\, dx$.

$\dashv$

Example 6.4 (Yor[**?**]). Let $X = \int_0^\infty \varphi(t)\, dW_t$, for some deterministic square–integrable φ. Note that, conditional on $\mathcal{F}_t$, the random variable X is Gaussian with mean $m_t = \int_0^\varphi(s)\, dW_s$ and variance $\sigma_t^2 = \int_t^\infty \varphi(s)^2\, ds$. To use Corollary 6.3, we must find $\rho(x,s)$ so that

$$\frac{1}{\sqrt{2\pi\sigma_t^2}}\, e^{-(x-m_t)^2/2\sigma_t^2} = \frac{1}{\sqrt{2\pi\sigma_0^2}}\, e^{-x^2/2\sigma_0^2}\, e^{\int_0^t \rho(x,s)\, dW_s - \frac{1}{2}\int_0^t \rho(x,s)^2\, ds}$$

Put $M_t = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-(x-m_t)^2/2\sigma_t^2}$. Some messy but straightforward calculations show that $dM_t = M_t \frac{x-m_t}{\sigma_t^2}\varphi(t)\, dW_t$, and thus that $M_t = e^{\int_0^t \rho(x,s)\, dW_s - \frac{1}{2}\int_0^t \rho(x,s)^2\, ds}$ for

$$\rho(x,t) = \frac{x-m_t}{\sigma_t^2}\varphi(t)$$

Under suitable integrability conditions (see the Remarks that follow), therefore, we see that W is a semimartingale in the enlargement $\mathbb{G}$ of $\mathbb{F}$ by X, with decomposition

$$W_t = \tilde{W}_t + \int_0^t \frac{\varphi(s)}{\sigma_s^2}\left(\int_s^\infty \varphi(u)\, dW_u\right)\, ds$$

where $\tilde{W}$ is a $\mathbb{G}$–Brownian motion.
In particular, if $\varphi(t) = I_{[0,T]}$, then $X = \int_0^\infty \varphi(t)\, dW_t = W_T$, and hence

$$W_t = \tilde{W}_t + \int_0^{t\wedge T} \frac{W_T - W_s}{T-s}\, ds$$

as before.

$\square$

Remarks 6.5 Note that if $\mathbb{G}$ is the enlargement of $\mathbb{F}$ by W_T, then Jacod's criterion $(\mathbf{J}')$ fails to hold on $[0,\infty)$. It does, however, hold on $[0,T)$. Indeed, W_T is clearly Gaussian conditional on $\mathcal{F}_t$ for $t < T$. Thus if $Q_t(\omega, dx)$ are regular conditional versions of $\mathbb{P}(W_T \in dx|\mathcal{F}_t)$, then $Q_t(\omega,dx)$ is absolutely continuous w.r.t Lebesgue measure. However, for $t \geq T$, $Q_t(\omega,dx)$ is the point mass $\delta_{W_T(\omega)}(dx)$, and it is impossible to find a single measure η such that $\delta_r(dx) << \eta(dx)$ for all $r \in \mathbb{R}$.

Jeulin and Yor[**?**] show that, in the Brownian framework, a $(\mathbb{F},\mathbb{P})$–local martingale M is a $(\mathbb{G},\mathbb{P})$–semimartingale (where $\mathbb{G}$ is the enlargement of $\mathbb{F}$ by W_T) iff $\int_0^T (1-s)^{\frac{1}{2}}\, |d[M,W]_s| < \infty$, in which case

$$M_t - \int_0^{t\wedge T} \frac{W_1 - W_s}{T-s}\, d[M,W]_s$$

is a $\mathbb{G}$–local martingale. Thus if $M = \int_0^\cdot m_s\, dW_s$, then M remains a $\mathbb{G}$–semimartingale iff $\int_0^T \frac{|m_s|}{\sqrt{T-s}}\, ds < \infty$.

Jeulin and Yor[?] further show that there exists a deterministic square–integrable m_s such that $\int_0^T \frac{|m_s|}{\sqrt{T-s}}\, ds = \infty$. It follows that not every $(\mathbb{F}, \mathbb{P})$–local martingale is a $(\mathbb{G}, \mathbb{P})$–semimartingale, i.e. that $(\mathbf{H}')$ fails.

□

We follow Protter[?] to elaborate on the results of Jeulin and Yor mentioned above:

Lemma 6.6 (Jeulin's Lemma) *On a general filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ satisfying the usual conditions, let $R : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ be a positive measurable process, with the properties that*

(i) all R_s are identically distributed, with common law μ satisfying $\mu\{0\} = 0$ and $\int_0^\infty x\, \mu(dx) < \infty$;

(ii) each R_s is independent of $\mathcal{F}_s$.

Suppose that A is a positive $\mathbb{F}$–predictable process with $\int_0^t A_s\, ds < \infty$ a.s. for each $t \geq 0$. Then the following sets are $\mathbb{P}$–a.s. equal:

$$\left\{\int_0^\infty R_s A_s\, ds < \infty\right\} = \left\{\int_0^\infty A_s\, ds\right\}$$

Proof: ([?], Jeulin[?]) We first show that $\left\{\int_0^\infty R_s A_s\, ds < \infty\right\} \subseteq \left\{\int_0^\infty A_s\, ds\right\}$. Let $E \in \mathcal{F}$ with $\mathbb{P}(E) > 0$, and let $J_t = \mathbb{E}[I_E|\mathcal{F}_t]$ be the càdlàg version of this martingale.

Let $J_* = \inf_t J_t$. Note that $\{J_* > 0\} \supseteq E$: For if $B \subseteq E$ is an event with $\mathbb{P}(B) > 0$, then

$$\mathbb{E}[J_t I_B] = \mathbb{E}[I_E \mathbb{E}[I_B|\mathcal{F}_t]] \geq \mathbb{E}[I_B \mathbb{E}[I_B|\mathcal{F}_t]] = \mathbb{E}[(\mathbb{E}[I_B|\mathcal{F}_t]^2] \geq (\mathbb{E}[I_B])^2 = \mathbb{P}(B)$$

where we used the fact that $\mathbb{E}\big[X\mathbb{E}[Y|\mathcal{F}_t]\big] = \mathbb{E}\big[\mathbb{E}[X|\mathcal{F}_t]\ \mathbb{E}[Y|\mathcal{F}_t]\big] = \mathbb{E}\big[\mathbb{E}[X|\mathcal{F}_t] Y\big]$ and also Jensen's inequality. Now let $\{q_n : n \in \mathbb{N}\}$ be an enumeration of the rationals. Then $\mathbb{E}[(\inf_{n \leq m} J_{q_n}) I_B] \geq \mathbb{P}(B)$ as well. It follows that $\mathbb{E}[J_* I_B] \geq \mathbb{P}(B)$ for all $B \subseteq E$, and thus that $J_* > 0$ a.s. on E.

Let F be the common distribution function of the R_t, i.e. $F(x) = \mu(0, x] = \mathbb{P}(R_t \leq x)$.

Note that

$$\mathbb{E}[I_E R_t|\mathcal{F}_t] = \mathbb{E}\left[I_E \int_0^\infty I_{\{R_t > u\}}\, du|\mathcal{F}_t\right] = \int_0^\infty \mathbb{E}[I_E I_{\{R_t > u\}}|\mathcal{F}_t]\, du$$

As $I_E I_{\{R_t > u\}} = (I_E - I_{\{R_t \leq u\}})^+$, we have $\mathbb{E}[I_E I_{\{R_t > u\}}|\mathcal{F}_t] \geq \mathbb{E}[I_E - I_{\{R_t \leq u\}}|\mathcal{F}_t]^+$, by Jensen's inequality. Using the fact that R_t is independent of $\mathcal{F}_t$, it now follows that

$$\mathbb{E}[I_E R_t|\mathcal{F}_t] \geq \int_0^\infty (J_t - F(u))^+\, du = \Phi(J_t)$$

where

$$\Phi(x) := \int_0^\infty (x - F(u))^+\, du$$

The function Φ is increasing and continuous on $[0,1]$. Furthermore, if $x > 0$, then $\Phi(x) > 0$ also, as $F(0) = \mu\{0\} = 0$ by assumption.

Now replace the arbitrary event E by the events

$$E_n = \left\{ \int_0^\infty R_t A_t \; dt \leq n \right\}$$

with $J_t^n = \mathbb{E}[I_{E_n} | \mathcal{F}_t]$ (càdlàg), to deduce that

$$\mathbb{E}\left[\int_0^\infty \Phi(J_t^n) A_t \; dt\right] \leq \mathbb{E}\left[\int_0^\infty \mathbb{E}[I_{E_n} R_t | \mathcal{F}_t] A_t \; dt\right] = \mathbb{E}\left[I_{E_n} \int_0^\infty R_t A_t \; dt\right] \leq n\mathbb{P}(E_n) < \infty$$

so that $\int_0^\infty \Phi(J_t^n) A_t \; dt < \infty$ a.s. Since

$$\Phi(J_*^n) \int_0^\infty A_t \; dt \leq \int_0 \Phi(J_t^n) A_t \; dt \qquad \text{a.s.}$$

we see that

$$\int_0^\infty A_t \; dt < \infty \qquad \text{a.s. on } \; E_n$$

Thus $\int_0^\infty A_t \; dt < \infty$ a.s. on $\bigcup_n E_n = \{\int_0^\infty R_t A_t \; dt < \infty\}$, i.e. $\{\int_0^\infty A_t \; dt < \infty\} \supseteq \{\int_0^\infty R_t A_t \; dt < \infty\}$ a.s.

We now show the reverse inclusion: Note that if T is a stopping time, then

$$\begin{aligned}
\mathbb{E}\left[\int_0^T R_s A_s \; ds\right] &= \int_0^\infty \mathbb{E}[I_{[0,T]} R_s A_s] \; ds \\
&= \int_0^\infty \mathbb{E}[I_{[0,T]} A_s \mathbb{E}[R_s | \mathcal{F}_{T\wedge s}] \; ds \\
&= \mathbb{E}\left[\int_0^T A_s \mathbb{E}\big[\mathbb{E}[R_s|\mathcal{F}_s] | \mathcal{F}_{T\wedge s}\big] \; ds\right] \\
&= \alpha \mathbb{E}\left[\int_0^T A_s \; ds\right]
\end{aligned}$$

where $\alpha = \mathbb{E}[R_s]$ is the common mean of the R_s, which is finite by assumption.

Define the stopping time $T_n = \inf\{t > 0 : \int_0^t A_s \; ds > n\}$, so that $\mathbb{E}[\int_0^{T_n} A_s \; ds] \leq n$. Now

$$\mathbb{E}\left[\int_0^{T_n} R_s A_s \; ds\right] = \alpha \mathbb{E}\left[\int_0^{T_n} A_s \; ds\right] \leq \alpha n \langle \infty$$

so that $\int_0^{T_n} R_s A_s \; ds < \infty$ a.s. If $\omega \in \{\int_0^\infty A_s \; ds < \infty\}$, then there exists n (depending on ω) such that $T_n(\omega) = \infty$, and so $\int_0^\infty R_s A_s \; ds(\omega) = \int_0^{T_n} R_s A_s \; ds(\omega) < \infty$, i.e. we also have $\omega \in \{\int_0^\infty R_s A_s \; ds < \infty\}$.

$\dashv$

Equipped with the lemma, we can prove the result of Jeulin–Yor[**?**]:

Theorem 6.7 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ support a Brownian motion W with natural filtration $\mathbb{F}$, and let $\mathbb{G}$ be the enlargement of $\mathbb{F}$ by W_T. Then an $(\mathbb{F}, \mathbb{P})$–local martingale is an $(\mathbb{G}, \mathbb{P})$–semimartingale iff $\int_0^T (T-s)^{-\frac{1}{2}} \, |d[M,W]|_s < \infty$. a.s.*
In that case

$$M_t - \int_0^{t\wedge T} \frac{W_T - W_s}{T-s} \, d[M,W]_s$$

is a $(\mathbb{G}, \mathbb{P})$–local martingale.

Proof: (Protter[**?**]) Write $M_t = M_0 + \int_0^t m_s \, dW_s$ by the martingale representation theorem, where m is predictable with $\int_0^t m_s^2 \, ds < \infty$ a.s. We know that W is a $(\mathbb{G}, \mathbb{P})$–semimartingale with canonical decomposition $W_t = \tilde{W}_t - \int_0^{t\wedge T} \frac{W_T - W_s}{T-s} \, ds$. Using Theorem 5.2, we see that M is a $(\mathbb{G}, \mathbb{P})$–semimartingale iff $\int_0^t |m_s| \frac{|W_T - W_s|}{T-s} \, ds < \infty$ a.s. for $0 \le t \le T$. Now apply Jeulin's lemma, with $A_t = \frac{|m_t|}{\sqrt{T-t}}$ and $R_t = I_{\{t<T\}} \frac{|W_T - W_t|}{\sqrt{T-t}}$. (Note that for $t < T$, we have R_t independent of $\mathcal{F}_t$, with law $N(0,1)$, so that Jeulin's Lemma is applicable.) Then $\int_0^t |m_s| \frac{|W_T - W_s|}{T-s} \, ds = \int_0^t A_s R_s \, ds$ which is finite a.s. iff $\int_0^t A_s \, ds = \int_0^t \frac{m_s}{\sqrt{1-s}} \, ds$ is finite a.s. Now by the associative law for integrals,

$$\int_0^t \frac{|m_s|}{\sqrt{1-s}} \, ds = \int_0^t \frac{1}{\sqrt{T-s}} \, |d[m \bullet W, W]_s| = \int_0^t \frac{1}{\sqrt{T-s}} \, |d[M,W]_s|$$

Again by Theorem 5.2, we see that $(m \bullet W)_t - \int_0^{T\wedge t} \frac{m_s(W_T - W_s)}{T-s} \, ds = M_t - \int_0^{T\wedge t} \frac{W_T - W_s}{T-s} \, d[M,W]_s$ is a $(\mathbb{G}, \mathbb{P})$–local martingale.

$\dashv$

Example 6.8 Here is an example of Jeulin and Yor[**?**] of a $(\mathbb{F}, \mathbb{P})$–local martingale which is not a $(\mathbb{G}, \mathbb{P})$–semimartingale (where $\mathbb{F}$ is the natural filtration of Brownian motion W, and $\mathbb{G}$ is the enlargement of $\mathbb{F}$ with W_1): Take an $\alpha \in (\frac{1}{2}, 1)$, and let $m_s = (1-s)^{-\frac{1}{2}}(-\ln(1-s))^{-\alpha} I_{\{\frac{1}{2}<s<1\}}$. Then m is a deterministic predictable function with $\int_0^1 m_s^2 \, ds < \infty$, so that $M = m \bullet W$ is defined. However, $\int_0^1 \frac{m_s}{\sqrt{1-s}} = \infty$, and so M is a $(\mathbb{F}, \mathbb{P})$–local martingale, but not a $(\mathbb{G}, \mathbb{P})$–semimartingale.

$\square$

A Monotone Class Theorems

The monotone class theorems are a collection of related results that prove that if a certain "nice" set of measurable functions satisfy a certain property, then all (bounded) measurable function have that property. In the literature, one frequently finds that a proof will verify a property for indicator functions, and then assert that "the rest follows by a monotone class argument".

The monotone class theorems are based on a kind of "decomposition" of a σ–algebra into a part that meshes nicely with the properties of measures (λ–systems) and part which doesn't, but which is nevertheless very simple, and meshes nicely with multiplication (π–systems):

Definition A.1 Let $\mathcal{C}$ be a collection of subsets of Ω

(a) $\mathcal{C}$ is called a π–system if it is closed under finite intersections.

(b) $\mathcal{C}$ is called a λ–system if

  (i) $\Omega \in \mathcal{C}$;

  (ii) $A, B \in \mathcal{C}$ and $A \subseteq B$ implies $B - A \in \mathcal{C}$;

  (iii) If $A_1, A_2, \cdots \in \mathcal{C}$ and $A_n \uparrow A$, then $A \in \mathcal{C}$.

□

We denote by $\pi(\mathcal{C})$ and $\lambda(\mathcal{C})$ the π–, respectively, λ–system *generated* by $\mathcal{C}$ The following lemma is easy to prove.

Lemma A.2 *A family $\mathcal{C}$ of subsets of Ω is a σ–algebra iff it is* both *a π–system and a λ–system.*

□

The following technical result often allows us to work with "easy" π–systems, instead of the "difficult" σ–algebras:

Theorem A.3 (Dynkin's Lemma, Monotone Class Theorem)

(a) If $\mathcal{C}$ is a π–system on Ω, then

$$\lambda(\mathcal{C}) = \sigma(\mathcal{C})$$

(b) Suppose that $\mathcal{C}$ is a π–system and that $\mathcal{D}$ is a λ–system (both on a set Ω), and also that $\mathcal{C} \subseteq \mathcal{D}$. Then $\sigma(\mathcal{C}) \subseteq \mathcal{D}$.

Proof: (a) Let $\mathcal{D} = \lambda(\mathcal{C})$. By Lemma A.2, it suffices to show that $\mathcal{D}$ is a π–system. We do this in two steps.
STEP I: Fix $C \in \mathcal{C}$, and define

$$\mathcal{D}_C = \{A \in \mathcal{D} : A \cap C \in \mathcal{D}\}$$

Then $\mathcal{C} \subseteq \mathcal{D}_C \subseteq \mathcal{D}$ (because $\mathcal{C}$ is a π–system). Then $\mathcal{D}_C$ is easily shown to be a λ–system containing $\mathcal{C}$, so that $\mathcal{D}_C = \mathcal{D}$.
STEP II: Now, fix any $D \in \mathcal{D}$, and define

$$\mathcal{D}^D = \{A \in \mathcal{D} : A \cap D \in \mathcal{D}\}$$

First note that if $C \in \mathcal{C}$, then $\mathcal{D}_C = \mathcal{D}$, so $D \in \mathcal{D}_C$. It follows that $D \cap C \in \mathcal{D}$, and thus that $C \in \mathcal{D}^D$, for every $C \in \mathcal{C}$. Thus $\mathcal{C} \subseteq \mathcal{D}^D$, for all $D \in \mathcal{D}$.
It follows as above that $\mathcal{D}^D$ is a λ–system, and thus that $\mathcal{D}^D = \mathcal{D}$, for all $D \in \mathcal{D}$.

In particular, if $A, B \in \mathcal{D}$, then $A \in \mathcal{D}^B$, and so $A \cap B \in \mathcal{D}$. This shows that $\mathcal{D}$ is a π–system, and thus a σ–algebra.

(b) follows directly from (a).

⊣

Definition A.4 (a) A collection $\mathcal{A}$ of bounded real–valued functions on a set Ω is called an *algebra*[1] if it is a vector space and closed under multiplication.

(b) A collection $\mathcal{H}$ of bounded real–valued functions on a set Ω is called a *monotone vector space* iff

(i) $\mathcal{H}$ is a vector space over $\mathbb{R}$.

(ii) The constant function 1 belongs to $\mathcal{H}$.

(iii) If $(f_n)_n$ is a uniformly bounded increasing sequence of non–negative members of $\mathcal{H}$, then $\lim_n f_n \in \mathcal{H}$.

□

Theorem A.5 (Monotone Class Theorem)
Let $\mathcal{H}$ be a monotone vector space on S. Let $\mathcal{A}$ be a π–system on S with the property that $I_A \in \mathcal{H}$ for every $A \in \mathcal{A}$.
Then every bounded $\sigma(\mathcal{A})$–measurable function belongs to $\mathcal{H}$.

Proof: Let $\mathcal{D} = \{F \subseteq S : I_F \in \mathcal{H}\}$. It is not hard to show that $\mathcal{D}$ is a λ–system. By Theorem A.3, $\mathcal{D} \supseteq \sigma(\mathcal{A})$.

Let h be a non–negative, bounded $\sigma(\mathcal{A})$–measurable function, with upper bound K, i.e.

$$0 \leq h(s) \leq K \qquad \text{for all } s \in S$$

If we define

$$h_n(s) = \sum_{k=1}^{K2^n} \frac{k-1}{2^n} I_{A(n,k)}(s) \qquad \text{where} \qquad A(n,k) = \left\{ s \in S : \frac{k-1}{2^n} \leq h(s) < \frac{k}{2^n} \right\}$$

then the h_n are simple functions with $h_n \uparrow h$. Since h is $\sigma(\mathcal{A})$–measurable, each $A(n,k) \in \mathcal{D}$, i.e. $I_{A(n,k)} \in \mathcal{H}$. Because $\mathcal{H}$ is a vector space, we now see that $h_n \in \mathcal{H}$ for each $n \in \mathbb{N}$. Thus $h \in \mathcal{H}$ as well.

We have now shown that every non–negative bounded $\sigma(\mathcal{A})$–measurable function belongs to $\mathcal{H}$. The same result can be obtained for arbitrary bounded h by splitting into positive and negative parts: $h = h^+ - h^-$.

$\dashv$

Here is another such result:

Theorem A.6 (Monotone Class Theorem) *Let $\mathcal{M}$ be a collection of bounded real–valued functions on Ω which is closed under multiplication. Suppose that $\mathcal{H}$ is a monotone vector space which is closed under uniform convergence, such that $\mathcal{H} \supseteq \mathcal{M}$. Then every bounded $\sigma(\mathcal{M})$–measurable function belongs to $\mathcal{H}$.*

[1] This is not to be confused with a family of sets closed under complementation and finite unions/intersections.

Proof: We sketch the proof given in Dellacherie and Meyer[**?**]. We may assume that $1 \in \mathcal{M}$. Let $\mathcal{A}$ be an algebra which is maximal such that $\mathcal{M} \subseteq \mathcal{A} \subseteq \mathcal{H}$ — such $\mathcal{A}$ exists by Zorn's Lemma. The function $x \mapsto |x|$ can be uniformly approximated by polynomials on every compact interval of $\mathbb{R}$. (To see this note that $|x| = (1-(1-x^2))^{\frac{1}{2}}$, and that the Taylor series of $z \mapsto (1-z)^{\frac{1}{2}}$ converges uniformly on $[-1.1]$.) Now $\mathcal{A}$ is closed under uniform convergence because $\mathcal{H}$ is, by the maximality of $\mathcal{A}$. It follows that if $f \in \mathcal{A}$, then also $|f| \in \mathcal{A}$. In particular, given $f, g \in \mathcal{A}$, we see that $f^{\pm} = \frac{|f| \pm f}{2} \in \mathcal{A}$, and hence that $f \vee g = g + (f-g)^+ \in \mathcal{A}$ and $f \wedge g = f + g - f \vee g \in \mathcal{A}$.

We claim that $\mathcal{A}$ is a monotone vector space. For suppose that g is the limit of a uniformly bounded increasing sequence $(g_n)_n$ of members of $\mathcal{A}$. It is easy to see that the algebra generated by $\mathcal{A} \cup \{g\} \subseteq \mathcal{H}$, so that $g \in \mathcal{A}$ by maximality of $\mathcal{A}$.

Let $\mathcal{C} = \{C \subseteq \Omega : I_C \in \mathcal{A}\}$. Since $\mathcal{A}$ is an algebra and $1 \in \mathcal{A}$, we see that $\mathcal{C}$ is a closed under finite intersections and complementation, and thus also under finite unions. Since $\mathcal{A}$ is monotone, $\mathcal{C}$ is closed under countable unions, and thus a σ–algebra.

Since every non–negative bounded measurable function can be uniformly approximated from below by simple measurable functions (as in the proof of Theorem A.5), we see that $\mathcal{A}$ contains all $\mathcal{C}$–measurable functions.

To complete the proof, we need only show that $\sigma(\mathcal{M}) \subseteq \mathcal{C}$, i.e. that $\{f \leq c\} \in \mathcal{C}$ for all $f \in \mathcal{M}$ and all $c \in \mathbb{R}$. Clearly it suffices to show that $\{f \geq 1\} \in \mathcal{C}$ for all $f \in \mathcal{A}$, i.e. that $I_{\{f \geq 1\}} \in \mathcal{A}$ for all $f \in \mathcal{A}$. Now if $f \in \mathcal{A}$, then $g = (f \wedge 1)^+ \in \mathcal{A}$. Now the sequence of n^{th} powers of g has $g^n \downarrow I_{\{f \geq 1\}}$, so that $I_{\{f \geq 1\}} \in \mathcal{A}$ because $\mathcal{A}$ is a monotone algebra.

$\dashv$

Remarks A.7 Protter([**?**], p.7) states that we may drop the assumption that $\mathcal{H}$ is closed under uniform convergence in Theorem A.6: A monotone vector space is always closed under uniform convergence.

$\square$

B Completions, the Usual Hypotheses, etc.

Here follow some definitions and remarks:

1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. $\mathcal{F}$ is said to be complete iff it contains all $\mathbb{P}$–negligible sets: Whenever $A \subseteq B$ and $B \in \mathcal{F}$ with $\mathbb{P}(B) = 0$, then $A \in \mathcal{F}$.

2. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, and that $\mathcal{G} \subseteq \mathcal{F}$. Then the completion $\mathcal{G}^{\mathbb{P}}$ of $\mathcal{G}$ in $\mathcal{F}$ is the σ–algebra with the following property:

$$A \in \mathcal{G}^{\mathbb{P}} \qquad \text{iff there exists } B \in \mathcal{F} \text{ such that} \qquad \mathbb{P}(A \Delta B) = 0$$

   Equivalently,

$$A \in \mathcal{G}^{\mathbb{P}} \quad \text{iff there exists } B \in \mathcal{G} \text{ and null sets } M, N \in \mathcal{F} \text{ such that } B - N \subseteq A \subseteq B \cup M$$

3. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is right–continuous if and only if

$$\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s \qquad \text{for all } t > 0$$

4. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to satisfy the *usual conditions* (w.r.t a probability measure $\mathbb{P}$) iff
   (i) $\mathbb{F}$ is right– continuous; and,
   (ii) $\mathcal{F}_0$ contains all the $\mathbb{P}$–null sets in $\mathcal{F}$.
5. If $\mathcal{F}$ is a filtration, we can augment it to satisfy the usual conditions, as follows: Let
$$\mathcal{G}_t = \bigcap_{s>t}(\mathcal{F}_s \cap \sigma(\mathcal{N}))$$
where $\mathcal{N}$ is the collection of all $\mathbb{P}$–null sets.
6. The imposition of the usual conditions is essential for the existence of regular versions of stochastic processes: Recall that a $(\mathbb{F}, \mathbb{P})$– submartingale $X = (X_t)_{t\geq 0}$has a càdlàg version iff the map $t \mapsto \mathbb{E}X_t$ is right–continuous. However, this requires the usual conditions; cf. Karatzas and Shreve[**?**]. To be precise, if $X = (X_t)_t$ is an $((\mathcal{F}_t)_t, \mathbb{P})$–submartingale (not assuming the usual conditions on $(\mathcal{F}_t)_t$), then $X^+ = (X_{t+})_t$ is a càdlàg $((\mathcal{F}_{t+})_t, \mathbb{P})$–submartingale. Here $X_{t+} = \lim_{s\downarrow t} X_s$ exists for all t, a.s. If $(\mathcal{F}_t)_t$ is right–continuous, then X^+ is adapted to $(\mathcal{F}_t)_t$, and can then be shown to be a modification of X. Cf. [**?**] p.16 for more details.

C Regular Conditional Probabilities

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(S, \mathcal{S}), (T, \mathcal{T})$ be measurable spaces. Given random elements $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{S})$ and $Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (T, \mathcal{T})$, we know how to define $\mathbb{P}(X \in A|Y)$ for $A \in \mathcal{S}$:
$$\mathbb{P}(X \in A|Y) = \mathbb{E}[I_A(X)|\sigma(Y)]$$
is just a version of a particular conditional expectation (which is a.s. unique).

We want, however, to make sense of the expression
$$\mathbb{P}(X \in A|Y = y) \qquad \text{for } A \in \mathcal{S}, y \in T$$
Consider the joint law $\mathbb{P}_{X,Y}$ on $(S \times T, \mathcal{S} \otimes \mathcal{T})$, given by
$$\mathbb{P}_{X,Y}(A \times B) = \mathbb{P}(X \in A, Y \in B)$$
We ought then be able to write
$$\mathbb{P}_{X,Y}(A \times B) = \int_B \mathbb{P}(X \in A|Y = y)\, \mathbb{P}_Y(dy) = \int_B \mathbb{P}_X^y(A)\, \mathbb{P}_Y(dy)$$
where $\mathbb{P}_Y$ is the law of Y on $(T, \mathcal{T})$, and $\mathbb{P}_X^y(A) = \mathbb{P}(X \in A|Y = y)$. If we can do this, we have *disintegrated* the joint law. However, to be able to do it, we clearly require that

- For $y \in T$, each map P_X^y is a probability measure on $(S, \mathcal{S})$, and
- For fixed $A \in \mathcal{S}$, the map $y \mapsto \mathbb{P}(X \in A|Y = y)$ is measurable, so that we can perform the integration.

Definition C.1 Given two measurable spaces $(S,\mathcal{S}),(T,\mathcal{T})$, a map $\mu : T\times\mathcal{S}\to\bar{\mathbb{R}}^+$ is called a *stochastic kernel* from $(T,\mathcal{T})$ to $(S,\mathcal{S})$ iff

(i) The map $t\mapsto\mu(A,t)$ is $\mathcal{T}$–measurable in $t\in T$ for fixed $A\in\mathcal{S}$, and

(ii) The map $A\mapsto\mu(A,t)$ is a probability measure on $(S,\mathcal{S})$ for fixed $t\in T$.

If $(T,\mathcal{T})=(\Omega,\mathcal{F})$ is a probability space, then the stochastic kernel μ is called a *random measure.*

$\square$

Definition C.2 Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. Given random elements $X:(\Omega,\mathcal{F},\mathbb{P})\to(S,\mathcal{S})$ and $Y:(\Omega,\mathcal{F},\mathbb{P})\to(T,\mathcal{T})$, a *regular conditional probability of* X *given* Y is a random measure of the form

$$\mu(Y,A)=\mathbb{P}[X\in A|Y]\quad\text{a.s.}$$

where μ is a stochastic kernel from T to S.
We then define $\mathbb{P}[X\in A|Y=y]:=\mu(y,A)$.

$\square$

Remarks C.3 Given $X:(\Omega,\mathcal{F},\mathbb{P})\mapsto(S,\mathcal{S})$ and a sub–σ–algebra$\mathcal{G}$ of $\mathcal{F}$, one often encounters the notion of a regular conditional probability of the form $\mathbb{P}[X\in\cdot|\mathcal{G}]$. By this is meant a version of $\mathbb{P}[X\in\cdot|\mathcal{G}]$ which is a stochastic kernel from $(\Omega,\mathcal{F})$ to $(S,\mathcal{S})$, i.e. a map $\nu:\Omega\times\mathcal{S}\to\bar{\mathbb{R}}^+$ having

(i) $A\mapsto\nu(\omega,A)$ is a.s. a probability measure on $(S,\mathcal{S})$;

(ii) $\omega\mapsto\nu(\omega,A)$ is $\mathcal{F}$–measurable — actually, $\mathcal{G}$–measurable, because in addition

(iii) $\nu(\cdot,A)$ is a version of $\mathbb{E}[X\in A|\mathcal{G}](\cdot)$.

The preceding Defn C.2 includes this as a special case: Put $Y=\mathrm{Id}_\Omega:(\Omega,\mathcal{F})\to(\Omega,\mathcal{G})$, and Then $\sigma(\mathrm{Id})=\mathcal{G}$, so that $\mu(\omega,A)=\mu(Y,A)(\omega)=\mathbb{P}[X\in A|\mathcal{G}]$ a.s.

$\square$

Regular conditional probabilities do not always exist — cf. Rogers and Williams I.43 for a counterexample — but some mild topological conditions on the state space $(S,\mathcal{S})$ will ensure existence: A *Borel space* is a measurable space $(S,\mathcal{S})$ with the property that there exists a Borel subset $B\subseteq[0,1]$ and a bijection $f:(S,\mathcal{S})\to(B,\mathcal{B}(B))$ such that both f,f^{-1} are measurable. In particular, it is known that any Polish space (equipped with its Borel algebra) is a Borel space (cf. Parthasarathy[**?**]). Consequently, any Borel subset of a Polish space, is a Borel space.

Theorem C.4 (Existence of Regular Conditional Distributions) *Let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space, and let* $(S,\mathcal{S}),(T,\mathcal{T})$ *be measurable spaces, and assume in addition that* $(S,\mathcal{S})$ *is a Borel space. Given random elements* $X:(\Omega,\mathcal{F},\mathbb{P})\to(S,\mathcal{S})$ *and* $Y:(\Omega,\mathcal{F},\mathbb{P})\to(T,\mathcal{T})$. *Then there exists a unique regular conditional distribution of* X *given* Y *(i.e. there exists probability kernel* $\mu:T\times\mathcal{S}\to\bar{\mathbb{R}}^+$ *such that* $\mathbb{P}[X\in\cdot|Y]=\mu(Y,\cdot)$ *a.s., and any two such kernels are* $\mathbb{P}_Y$*–a.e. equal.)*

Proof: (Taken from Kallenberg[**?**].) Since S is a Borel space, we may assume w.l.o.g. that $S \in \mathcal{B}(\mathbb{R})$, and thus that X is real–valued, and $\mathcal{S} = S \cap \mathcal{B}(\mathbb{R})$.

Choose, for $q \in \mathbb{Q}$, a measurable function $f_q : T \to [0,1]$ such that

$$f_q(Y) = \mathbb{P}[X \leq q|Y] \text{ a.s.}$$

(this is possible, by the Doob–Dynkin Lemma) and define $f(t,q) = f_q(t)$. Let

$$T' = \{t \in T : f(t,q) \text{ is increasing in } q, \lim_{q\to\infty} f(t,q) = 1, \lim_{q\to-\infty} f(t,q) = 0\}$$

Note that $T' \in \mathcal{T}$ (because $T' = \bigcap_{q_1 \leq q_2}\{t : f_{q_1}(t) \leq f_{q_2}(t)\} \cap \bigcap_n \bigcup_q \{t : f_q(t) < \frac{1}{n}\} \cap \bigcap_n \bigcup_q \{t : f_q(t) > 1 - \frac{1}{n}\}$). Also note that each of the conditions defining T' holds at $Y(\omega)$ a.s., so that $Y \in T'$ a.s.

We now define $F : T \times \mathbb{R} \to [0,1]$ by

$$F(t,x) = I_{T'} \inf_{q>x} f(t,q) + I_{T'^c} I_{\{x \geq 0\}}$$

which has the property that $F(t,\cdot)$ is a distribution function on $\mathbb{R}$ for all $t \in \mathbb{T}$ (recall that a function is a distribution function precisely when it is right–continuous, tends to 1 at $+\infty$ and to 0 at $-\infty$). Let m_t be the Lebesgue–Stieltjes measure associated with $F(t,\cdot)$ (i.e. the unique probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ satisfying $m_t(-\infty,x] = F(t,x)$). Note also that, for fixed x, the map $t \mapsto F(t,x)$ is measurable in $t \in T$ (because, e.g. for $0 < u < 1$ and $x \geq 0$, we have $\{t : F(t,x) < u\} = \{t \in T' : \exists q \in \mathbb{Q}(q > x \wedge f(t,q) < u)\}$, etc.) So $m(t,B) = m_t(B)$ behaves like a stochastic kernel on the sets $B = (-\infty, x]$, and a monotone class argument shows that m is a stochastic kernel from $(T,\mathcal{T})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Furthermore, since $m(t,(-\infty,x]) = F(t,x)$, we have $m(Y,(-\infty,x]) = F(Y,x)$. Since also $f(Y,q) = \mathbb{P}[X \leq q|Y]$ a.s., the monotone convergence theorem for conditional expectations ensures that $F(Y,x) = \mathbb{P}[X \leq x|Y]$ a.s., so that $m(Y,(-\infty x]) = \mathbb{P}[X \leq x|Y]$ a.s. Another application of a monotone class theorem shows that

$$m(Y,B) = \mathbb{P}[X \in B|Y] \text{ a.s.} \qquad \text{for all } B \in \mathcal{B}(\mathbb{R})$$

The kernel m is almost what we seek: it is a kernel form T to $\searrow$, whereas we need a kernel from T to S. Note, however, that $m(Y,S^c) = 0$ a.s. (because X takes values in S). Define $\mu : T \times \mathcal{S} \to [0,1]$ by

$$\mu(t,\cdot) = \begin{cases} m(t,\cdot) & \text{if } m(t,S) = 1 \\ \delta_{s_0} & \text{else} \end{cases}$$

where $s_0 \in S$ is arbitrary. It is not hard to see that μ is a regular conditional distribution.

Uniqueness is easy: If μ' is another regular conditional distribution of X given Y, then since two versions of conditional expectation are a.s. equal, we have

$$\mu(Y,(-\infty,q]) = \mathbb{P}[X \leq q|Y] = \mu'(Y,(-\infty,q]) \text{ a.s.}$$

By a monotone class theorem, $\mu(Y,\cdot) = \mu'(Y,\cdot)$.

$\dashv$

Theorem C.5 (Disintegration) *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(S, \mathcal{S}), (T, \mathcal{T})$ be measurable spaces. Suppose we are given a sub–σ–algebra $\mathcal{G}$ of $\mathcal{F}$ and a random element X of S such that $\mathbb{P}[X \in \cdot|\mathcal{G}]$ has a regular conditional version ν. Also, consider a $\mathcal{G}$–measurable random element Y of T and a measurable function $f : S \times T \to \mathbb{R}$ with $\mathbb{E}|f(X,Y)| < \infty$. Then*

$$\mathbb{E}[f(X,Y)|\mathcal{G}] = \int_S f(s,Y)\ \nu(ds) \quad a.s. \tag{†}$$

Proof: (Taken from Kallenberg[**?**].) If $f = I_{A\times B}$, where $A \in \mathcal{S}, B \in \mathcal{T}$, then $\mathbb{P}[X \in A, Y \in B] = \mathbb{E}[I_B(Y)\mathbb{E}[I_A(X)|\mathcal{G}]\ ] = \mathbb{E}[I_B(Y)\int_S I_A(s)\ \nu(ds)] = \mathbb{E}\left[\int_S f(s,Y)\ \nu(ds)\right]$. Thus

$$\mathbb{E}[f(X,Y)] = \mathbb{E}\left[\int_S f(s,Y)\ \nu(ds)\right] \tag{$*$}$$

for $f = I_{A\times B}$. By a monotone class theorem, $(*)$ holds for all measurable indicator functions, then extends by linearity and the monotone convergence theorem to all non–negative measurable functions.

Now assume that $f \geq 0$ is measurable with $\mathbb{E}[f(X,Y) < \infty$, and fix $G \in \mathcal{G}$. Then (Y, I_A) is a $\mathcal{G}$–measurable random element of $T \times \{0,1\}$, so using $(*)$ we see that

$$\mathbb{E}[f(X,Y)I_A] = \mathbb{E}\left[\int_S f(s,Y)I_A\ \nu(ds)\right] = \mathbb{E}\left[\int_S f(s,Y)\ \nu(ds)\ I_A\right]$$

from which it follows that $\int_S f(s,Y)\ \nu(ds)$ is a version of $\mathbb{E}[f(X,Y)|\mathcal{G}]$. This proves the result for $f \geq 0$. The general result follows by decomposing a measurable f into a difference of its positive and negative parts.

$\dashv$

Taking $\mathcal{G} = \sigma(Y)$ in (†) in Thm. C.5, and $\nu(\omega, ds) = \mu(Y(\omega), ds)$ for we see that

Corollary C.6 *If $\mathbb{P}[X \in \cdot|Y]$ has a regular conditional distribution μ, then*

$$\mathbb{E}[f(X,Y)|Y] = \int_S f(s,Y)\ \mu(Y,ds)$$

whenever f is measurable with $\mathbb{E}|f(X,Y)| < \infty$.

□

Applying the expectation operator to both sides of the equation in the preceding corollary, we see that

Corollary C.7 *If $\mathbb{P}[X \in \cdot|Y]$ has a regular conditional distribution μ, then*

$$\mathbb{E}[f(X,Y)] = \mathbb{E}\left[\int_S f(s,Y)\ \mu(Y,ds)\right]$$

whenever f is measurable with $\mathbb{E}|f(X,Y)| < \infty$.

□

We end this section with a result on "nice", jointly measurable, densities. Recall that a σ–algebra is said to be *separable* if it is generated by a countable family of sets (i.e. $\mathcal{F}$ is separable iff $\mathcal{F} = \sigma(F_n : n \in \mathbb{N})$). The Borel algebra of a separable metrizable space X is clearly separable: For if D is a countable dense subset of X, then $\mathcal{B}(X) = \sigma(B(d, \frac{1}{n}) : d \in D, n \in \mathbb{N})$.

Theorem C.8 (Doob Theorem on Disintegration) *Let $\mu : \Omega \times \mathcal{S} \to \bar{\mathbb{R}}^+$ be a stochastic kernel from $(T, \mathcal{T})$ to $(S, \mathcal{S})$, and let η be a finite measure on $(S, \mathcal{S})$. Suppose in addition that:*

(i) $\mathcal{S}$ is separable.

(ii) Each measure $\mu(t, \cdot)$ is absolutely continuous w.r.t. η on $(S, \mathcal{S})$.

(iii) Each measure $\mu(t, \cdot)$ is finite on $(S, \mathcal{S})$, for all $t \in T$.

Then there is a non–negative $\mathcal{T} \otimes \mathcal{S}$–measurable function $g(t, s)$ such that

$$\mu(t, ds) = g(t, s)\ \eta(ds)$$

Proof: (Taken from Yor and Meyer[**?**]) Without loss of generality, we may assume that $\eta = \mathbb{P}$ is a probability measure. As $\mathcal{S}$ is separable, we may write $\mathcal{S} = \bigvee_n \mathcal{S}_n$, where the $\mathcal{S}_n$ form a sequence of sub–σ–algebras of $\mathcal{S}$ generated by finer and finer finite partitions $\mathcal{P}_n$. Define

$$g_n(t, s) = \frac{d\mu(t, ds)}{d\mathbb{P}(ds)}\Big|_{\mathcal{S}_n}$$

i.e. if $s \in A$, where $A \in \mathcal{P}_n$ is a block of the partition that generates $\mathcal{S}_n$, then $g^n(t, s) = \mu(t, A)/\mathbb{P}(A)$. That $g_n(t, s)$ is $\mathcal{T} \otimes \mathcal{S}_n$–measurable now follows from the measurability of $t \mapsto \mu(t, A)$:

$$\{g_n \le c\} = \bigcup_{A \in \mathcal{P}_n} \mu(\cdot, A)^{-1}(-\infty, c\mathbb{P}(A)] \times A \in \mathcal{T} \otimes \mathcal{S}_n$$

Now clearly $\frac{d\mu(t,\cdot)}{d\mathbb{P}}|_{\mathcal{S}_n} = \mathbb{E}_{\mathbb{P}}[\frac{d\mu(t,\cdot)}{d\mathbb{P}}|\mathcal{S}_n]$ a.s. , so that $(g_n(t, \cdot))_n$ is a uniformly integrable $(\mathbb{P}, (\mathcal{S}_n)_n)$–martingale for each t. Hence there is $g(t, s)$ such that

$$g_n(t, s) \to g(t, s) \qquad \mathbb{P}\text{–a.s. and in } L^1(\mathbb{P}), \text{ for each } t \in T$$

$\dashv$